PTO/SB/08a (08-03)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Not for submission under 37 CFR 1.99)

Application Number	10758370
Filing Date	2004-01-15
First Named Inventor	David M. Bargeron
Art Unit	2171
Examiner Name	unknown
Attorney Docket Number	MS306435.01/MSFTP504US

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	11095393			David Bargeron, et al.	
	2	11165070			David Bargeron	
	3	11171064			David Bargeron	

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20040205542		2004-10-14	David M. Bargeron, et al.	
	2	20040205545		2004-10-14	David M. Bargeron, et al.	
	3	20040252888		2004-12-16	David M. Bargeron, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	10758370
	Filing Date	2004-01-15
	First Named Inventor	David M. Bargeron
	Art Unit	2171
	Examiner Name	unknown
	Attorney Docket Number	MS306435.01/MSFTP504US

	4	20060050969		2006-03-09	Michael Shilman, et al.	
	5	20060045337		2006-03-02	Michael Shilman, et al.	

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] ¡	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1	VINAJAK R. BORKAR, et al., Automatically extracting structure from free text addresses, 2000, 6 pages, In Bulletin of the IEEE Computer Society Technical committee on Data Engineering. IEEE.	☐
	2	REMCO BOUCKAERT, Low level information extraction: A bayesian network based approach, 2002, 9 pages, In Proceedings of TextML 2002, Sydney, Australia.	☐
	3	CLAIRE CARDIE, et al., Proposal for an interactive environment for information extraction, 1998, 12 pages, Technical Report TR98-1702, 2.	☐
	4	RICH CARUANA, et al., High precision information extraction, August 2000, 7 pages, in KDD-2000 Workshop on Text Mining.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	10758370
	Filing Date	2004-01-15
	First Named Inventor	David M. Bargeron
	Art Unit	2171
	Examiner Name	unknown
	Attorney Docket Number	MS306435.01/MSFTP504US

	5	M. COLLINS, Discriminative training methods for hidden markov models : Theory and experiments with perception algorithms, July 2002, pages 1-8, In Proceedings of Empirical Methods in Natural Language Processing (EMNLP02).	☐
	6	CORINNA CORTES, et al., Support-vector networks. Machine Learning, 1995, 20(3): 273-297.	☐
	7	Y. FREUND, et al., Large margin classification using the perceptron algorithm, Machine earning, 37(3):277-296.	☐
	8	Y. FREUND, et al., Experiments with a new boosting algorithm, 1996, In International Conference on Machine Learning, pages 148-156.	☐
	9	T. KRISTJANSSON, et al., Interactive information extraction with constrained conditional random fields, 2004, In Proceedings of the 19th international conference on artificial intelligence, AAAI. pages, 412-418.	☐
	10	JOHN LAFFERTY, et al., Conditional random fields: Probabilistic models for segmenting and labeling sequence data, 2001, In Proc. 18th International Conf. on Machine Learning, pages 282-289. Morgan Kaufmann, San Francisco, CA.	☐
	11	M. MARCUS, et al., The penn treebank: Annotating predicate argument structure, 1994, pages 114-119.	☐
	12	ANDREW MCCALLUM, Efficiently inducing features of conditional random fields, 2003, 8 pages, In Nineteenth Conference on Uncertainty in Artificial Intelligence (UAI03).	☐
	13	ANDREW MCCALLUM, et al., Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons, 2003, 4 pages, In Hearst/Ostendorf, Eds, HLT-NAACL, Ass'n for Computational Linguistics, Edmonton, Alberta, Canada.	☐
	14	KAMAL NIGAM, et al., Using maximum entropy for text classification, 1999, 7 pages, In Proceedings of the IJCAI'99 Workshop on Information Filtering.	☐
	15	DAVID PINTO, et al., Table extraction using conditional random fields, 2003, 8 pages, In Proceedings of the ACM SIGIR'03, July 28-August 1, 2003, Toronto, Canada.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	10758370
	Filing Date	2004-01-15
	First Named Inventor	David M. Bargeron
	Art Unit	2171
	Examiner Name	unknown
	Attorney Docket Number	MS306435.01/MSFTP504US

	16	L.R. RABINER, A tutorial on hidden markov models and selected applications in speech recognition, 1989, In Proceedings of the IEEE, Volume 77, pages 257-286.	☐
	17	FEI SHA, et al., Shallow parsing with conditional random fields. In Hearst/Ostendorf, Eds, 2003, HLT-NAACL: Main Proceedings, pages 213-220, Ass'n for Computational Linguistics, Edmonton, Alberta, Canada.	☐
	18	J. STYLOS, et al., Citrine:providing intelligent copy-and-paste, 2005, In Proceedings of ACM Symposium on User Interface Software and Technology (UIST 2004), pages 185-188.	☐
	19	B. TASKAR, et al., Max-margin parsing, 2004, 8 pages, In Empirical Methods in Natural Language Processing (EMNLP04).	☐
	20	S. MAO, et al., Document structure analysis algorithms: A literature survey, January 2003, Vol. 5010, pp. 197-207, In Proc. SPIE Electronic Imaging.	☐
	21	M. KRISHNAMOORTHY, et al., Syntactic segmentation and labeling of digitized pages from technical journals, 1993, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, pp. 737-747.	☐
	22	J. KIM, et al., Automated labeling in document images, January 2001, page 1-12, In Document Recognition and Retrieval VIII, Vol. 4307. Available online at http://archive.nlm.nih.gov/pubs/kim/spie2001/spie2001.pdf, last checked April 2, 2006.	☐
	23	D. NIYOGI, et al., Knowledge-based derivation of document logical structure, 1995, page 472-475, In Third International Conference on Document Analysis and Recognition, Montreal, Canada.	☐
	24	A. CONWAY, Page Grammars and Page Parsing: A Syntactic Approach to Document Layout Recognition, 1993, In Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba Science City, Japan, pages 761-764.	☐
	25	E.G. MILLER, et al., Ambiguity and constraint in mathematical expression recognition, 1998, 8 pages, In Proceedings of the National Conference of Artificial Intelligence. American Association of Artificial Intelligence.	☐
	26	T. TOKUYASU, et al., Turbo recognition: a statistical approach to layout analysis, 2001, in Proceedings of the SPIE, San Jose, CA, Vol. 4307, pages 123-129.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	10758370
	Filing Date	2004-01-15
	First Named Inventor	David M. Bargeron
	Art Unit	2171
	Examiner Name	unknown
	Attorney Docket Number	MS306435.01/MSFTP504US

	27	T. KANUNGO, et al., Stochastic language model for style-directed physical layout analysis of documents, 2003, pages 583-596, In IEEE Transactions on Image Processing, Vol. 5, No. 5.	☐
	28	D. BLOSTEIN, et al., Applying compiler techniques to diagram recognition, In Proceedings of the 16th International Conference on Pattern Recognition, 2002, Vol. 3, pages 123-136.	☐
	29	J. F. HULL, Recognition of mathematics using a two dimensional trainable context-free grammar, Master's thesis, MIT, June 1996, 101 pages.	☐
	30	N. MATSAKIS, Recognition of handwritten mathematical expressions, May 1999, page 1-59, Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.	☐
	31	J. LAFFERTY, et al., Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, 2001, In Proceedings of the 18th International Conference on Machine Learning, Morgan Kaufmann, San Francisco, CA, pages 282-289.	☐
	32	E. CHARNIAK, et al., Edge-Based Best-First Chart Parsing, 1998, In Proceedings of the 14th National Conference on Artificial Intelligence, pages 127-133.	☐
	33	D. KLEIN, et al., A* parsing: Fast Exact Viterbi Parse Selection, Stanford University, 2001, 8 pages, Tech. Rep. dbpubs/2002-16.	☐
	34	Y. FREUND, et al., A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting, 1995, In Computational Learning Theory: Eurocolt '95, Springer-Verlag, pages 23-37.	☐
	35	I. PHILIPS, et al., Cd-rom Document Database Standard, In Proceedings of the 2nd International Conference on Document Analysis and Recognition (ICDAR), 1993, pages 478-483.	☐
	36	P. VIOLA, et al., Rapid Object Detection Using a Boosted Cascade of Simple Features, 2001, page 1-9, In Proceedings of the IEEE Conference on Computer Vision and Pattern	☐
	37	T. BREUEL, High Performance Document Layout Analysis, 10 pages, In 2003 Symposium on Document Image Understanding Technology, Greenbelt Maryland.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	10758370
	Filing Date	2004-01-15
	First Named Inventor	David M. Bargeron
	Art Unit	2171
	Examiner Name	unknown
	Attorney Docket Number	MS306435.01/MSFTP504US

	38	R. ZANIBBI, et al., A Survey of Table Recognition: Models, Observations, Transformations, and Inferences, International Journal of Document Analysis and Recognition, 2004, Vol. 7, No. 1. pages 1-16.	☐
	39	K. F. CHAN, et al., Mathematical Expression Recognition: A Survey, 2000, International Journal on Document Analysis and Recognition, Vol. 3, pages 3-15.	☐
	40	E. CHARNIAK, Statistical Techniques for Natural Language Parsing, AI Magazine, 1997, Vol. 18, No. 4, pages 33-44.	☐
	41	M. KAY, Chart Generation, In Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL '96), Santa Cruz, California, 1996, pages 200-204.	☐
	42	M. VISWANATHAN, et al., Document Recognition: An Attribute Grammar Approach, March 1996, In Proc. SPIE Vol. 2660, Document Recognition III, Vincent/Hull, Eds., pages 101-111.	☐
	43	C.D. MANNING, et al., Foundations of Statistical Natural Language Processing. The MIT Press, 1999, page 1-3.	☐
	44	TOBIAS SCHEFER, et al., Active Hidden Markov Models For Information Extraction, In Advances in Intelligent Data Analysis, 4th International Conference, IDA 2001, pages 309-318.	☐
	45	P. CHOU, Recognition Of Equations Using a 2-D Stochastic Context-Free Grammar, In SPIE Conference on Visual Communications and Image Processing, Philadelphia, PA, 1989, pages 852-863.	☐
	46	M. KAY, Algorithm Schemata And Data Structures In Syntactic Processing, 1986, pages 35-70.	☐
	47	MICHAEL SHILMAN, et al., Spatial Recognition and Grouping of Text and Graphics, Eurographics Workshop on Sketch-Based Interfaces and Modeling, 2004, 5 pages, Hughes/Jorge, Eds.	☐
	48	MICHAEL SHILMAN, et al., Recognition and Grouping of Handwritten Text in Diagrams and Equations, IWFHR 2004, September 2004, page 69-77, Toyko, Japan.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	10758370
	Filing Date	2004-01-15
	First Named Inventor	David M. Bargeron
	Art Unit	2171
	Examiner Name	unknown
	Attorney Docket Number	MS306435.01/MSFTP504US

	49	MICHAEL SHILMAN, et al., Recognizing Freeform Digital Ink Annotations, IAPR International Workshop on Document Analysis Systems, September 8-10, 2004, 12 pages, Florence, Italy.	☐
	50	MICHAEL COLLINS, et al., "Logistic Regression, AdaBoost, and Bregman Distances", Machine Learning, 48(1/2/3) 2002	☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Not for submission under 37 CFR 1.99)

Application Number	10758370
Filing Date	2004-01-15
First Named Inventor	David M. Bargeron
Art Unit	2171
Examiner Name	unknown
Attorney Docket Number	MS306435.01/MSFTP504US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☑ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Himanshu S. Amin/	Date (YYYY-MM-DD)	2006-08-22
Name/Print	Himanshu S. Amin	Registration Number	40894

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.